NICHE OVERLAP MEASURES AND HYPOTHESIS TESTING: A REVIEW WITH PARTICULAR REFERENCE TO EMPIRICAL APPLICATIONS

Cornelis R. Neet, Institut de Zoologie et d’Ecologie Animale, Université de Lausanne, Bâtiment de Biologie, CH-1015 Lausanne, Switzerland

Keywords. Niche overlap, Similarity, Hypothesis testing, Community organization, Competition.

Abstract. Niche overlap measures in single and multiple resource states, and the current status of hypothesis testing using niche overlap measures, are reviewed. Two main hypotheses are recognized. To test the hypothesis of niche identity, the comparison of replicated samples of intra- and interspecific niche overlap measures is suggested. To test the influence of interspecific interactions on niche overlap, the shufled-sticks model appears promising. Relying on the available information on the properties, bias and statistical parameters of niche overlap measures, some practical recommendations are made. Although single resource niche overlap measures have been extensively studied and may be applied to empirical data, comparisons between communities and measures in multiple resource states are more problematic and should be handled with caution. The present status of the niche overlap concept is discussed.

Introduction

The purpose of measuring niche overlap is either to quantify the degree that two species share a set of common resources in a common habitat, or, more generally, to quantify the degree of similarity in resource use between species. At least 24 different measures have been proposed (Table 1). These can be described either as distance measures, similarity indexes, correlation coefficients or statistical tests (Ludwig and Reynolds 1988). Reviews of the most commonly used measures are given by Hurlbert (1978), Smith and Zaret (1982), Zaret and Smith (1984) and Ludwig and Reynolds (1988).

Many authors, e.g. Colwell and Futuyma (1971), Sale (1974), Hurlbert (1978), Abrams (1980), Lawlor (1980a) and Case (1984), have discussed the measurement of niche overlap in connection with competition, i.e. niche overlap as an estimation of the interspecific competition coefficients (αij) of the Lotka-Volterra model on competition between two species (see also Ayala et al. 1973 for a review of the different models of the Lotka-Volterra family). However, most authors recognize that it is not possible to estimate competition coefficients from niche overlap measures, and some have seriously argued against this view, resulting in what seems to be a growing agreement that the relations between niche overlap and interspecific competition coefficients are at least complex (Planka 1976; Abrams 1980; Lawlor 1980a), and that niche overlap is neither necessary nor sufficient to infer competition (Schoener 1974; Cohen 1978; Holt 1987). Therefore, unless a specific symbol is available, the symbol Oij is preferable to αij for niche overlap measures.

Although niche overlap measures may not serve to evaluate competition coefficients, they may permit insight into the structure of communities, for example, by making the “niche overlap hypothesis” (Planka 1972) testable. This hypothesis predicts that maximal tolerable overlap should be lower in intensely competitive situations than in environments with lower demand/supply ratios (Planka 1976), and is a subject of much current research and debate (see Giller 1984, for a review). Another application is the use of the measure of niche overlap to test hypotheses one can make on ecological similarity between species within given communities, i.e. as MacArthur (1968) pointed out, to make testable statements about differences between niches.

In this paper, my aim is to review the available niche overlap measures, to review the techniques that have been suggested for hypothesis testing, and to select the most useful procedures for empirical field situations, where qualities such as simplicity of computation and precision of the decision rules may be expected to be of particular importance. Throughout this review, I will keep the definition of the niche adopted, among others, by Colwell and Futuyma (1971) and Hurlbert (1981), i.e. the realized niche as a set of resource states used by a given species. In this definition, a given resource state is equivalent to a given niche macrodimension, sensu Slobodchikoff and Schulz (1980).

Measures of niche overlap

The measures of niche overlap are listed in Table 1. Niche overlap has been taken in a broad sense, some

Table 1. Niche overlap measures, *lato sensu*. Column (r) indicates the number of resource states for which the measure is designed, column (a) indicates whether the measure incorporates estimates of resource availability. The last column includes all papers dealing with topics like variance estimators, statistical testing and relative bias of niche overlap measures.

<table>
<thead>
<tr>
<th>Author</th>
<th>r</th>
<th>a</th>
<th>parameter estimates, bias and statistical tests (authors)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(6) Lloyd (1967)</td>
<td>1</td>
<td>no</td>
<td>Maurer (1982).</td>
</tr>
<tr>
<td>(8) Levins (1968), $\alpha_{ij}$</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(9) Levins (1968), $d_{ij}$ (euclid. distance)</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(10) Terborgh and Diamond (1970)</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(11) May and MacArthur (1972), d/w</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(12) Pielou (1972)</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(13) Goodall (1973)</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(14) Pianka (1973)</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(15) Cody (1974), summation $O_{ij}$</td>
<td>k</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(16) Cody (1974), product $O_{ij}$</td>
<td>k</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(17) Schoener (1974)</td>
<td>1</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>(20) Harner and Whitmore (1977), $\beta_{ij}$</td>
<td>k</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(22) Petrakis (1979), $\Phi_{ij}$</td>
<td>1</td>
<td>no</td>
<td>Petrakis (1979).</td>
</tr>
<tr>
<td>(23) Petrakis (1979), $G$</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>(24) Case (1984), $\alpha_{ij}$</td>
<td>1</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>
measures having not been originally defined as niche overlap measures in the way we did here, but rather as competition coefficient estimates (Schoener 1974), habitat utilization similarity measures (Pielou 1972), or species distributional similarity measures (Goodall 1973). Some measures have been re-invented several times, like Renkonen’s proportional overlap index (Hurlbert 1978) that is often referred to as Schoener’s index (Schoener 1970), while other measures have undergone several improvements (e.g. MacArthur and Levins 1967, MacArthur 1972, May and MacArthur 1972, May 1973). In such cases, only the first author is mentioned.

Some niche overlap measures have also been suggested as niche breadth measures (overlap between resource use and resource availability), e.g. by Feinsinger et al. (1981) and Smith (1982). Petratis (1981, 1983) has discussed graphical methods that may help to establish relations between niche breadth and overlap measures.

**Niche overlap in a single resource state**

Most niche overlap measures are measures of similarity in resource use on one unique resource dimension (Table 1), based upon resource utilization data, e.g. the amounts of total resources consumed by each consumer species in each of a number of resource categories (Abrams 1980). Some measures apply only to continuous variables that are approximately normally or log-normally distributed (e.g. MacArthur and Levins 1967; MacArthur 1972; May and MacArthur 1972; May 1973; Maurer 1982; Case 1984). Although these measures are not primarily intended for use in empirical studies (Hurlbert 1978), they may be useful in particular cases (Southwood 1978; May 1973; Case 1984). However, in most situations, niche overlap measures have to deal with resource states that are subdivided into categories, i.e. a nominal scale (Hurlbert 1978). Among the single dimension measures that may be applied to nominal scales, some stand out as asymmetric measures, i.e. the overlap of species i on species j may differ from the overlap of species j on species i, e.g. Levins (1968) and Petratis (1979). This property is appealing when competition coefficients are to be estimated by niche overlap. However, it is a source of complication when the generalization of the measures to multiple resource dimensions is attempted (Slobodchikoff and Schulz 1980).

Only two single dimension measures take resource availability into account, those of Schoener (1974) and Hurlbert (1978). The first was designed as a competition coefficient estimator, and, as such, has recently been shown to be useful (Spiller 1986). Nevertheless, as a niche overlap measure, Schoener’s measure is too complex to be applicable in most situations since it requires the field estimation of consumption rates per unit time as well as other parameters, and, thus, requires exacting field work (Spiller 1986). On the other hand, Hurlbert’s measure is easier to measure, since, apart from resource utilization data, only relative frequencies of resource availability are necessary (Hurlbert 1978). This measure has, moreover, the interesting property of being not only an overlap measure, but also a selectivity (electivity) measure, as it varies between 0, for no overlap, through 1, for overlap with relative resource category use exactly proportional to relative resource category availability, to > 1, for niche overlap with convergence in resource selectivity. As such, this measure is especially useful. As Hurlbert (1978), Abrams (1982) and Hurlbert (1982) underlined, niche overlap measures including resource availabilities have the drawback of being limited to species coexisting on a common resource spectrum. They cannot be used for comparisons between communities which may differ in resource availability.

The remaining single dimension niche overlap measures are rather homogeneous in their properties except for their bias, i.e. their accuracy. The bias is usually evaluated by means of analytic or simulation techniques (e.g. Smith and Zaret 1982).

Bias has been evaluated for the measures of Renkonen (1938), Matusita (1955), Morisita (1959), Horn (1966), Van Beele and Ahmad (1974) and Petratis (1979) by Smith and Zaret (1982); for the measures of Renkonen (1938), Morisita (1959), Horn (1966) and Pianka (1973) by Linton et al. (1981); and for the measures of Renkonen (1938), Morisita (1959), Horn (1966) and Levins (1968) by Ricklefs and Lau (1980). Smith and Zaret (1982) show that the measure of Renkonen (1938) and the Ca measure of Horn (1966) are significantly biased and that all measures are increasingly biased as the number of resource categories rises. According to their study, the less biased measures are those of Matusita (1955) and Morisita (1959). Linton et al. (1981) found the measures they compared rather similar in efficiency, the Renkonen measure being the better (however, for technical reasons, some of their results should, according to Hurlbert (1982), be used with caution). Ricklefs and Lau (1980) show bias for the four measures they examined when sample sizes are small and overlap values close to 1 (complete overlap).

**Niche overlap in multiple resource states**

Only four measures make an explicit attempt to encompass several resource states, the two measures of Cody (1974) and the two of Harmer and Whitmore (1977). The measures suggested by Cody (1974) consist in the calculation of “general overlap” either by the sum or by the product of component overlaps. Cody calculated the overlaps in habitat use, in vertical foraging zones and in feeding behavior and then combined these three component overlaps to obtain a generalized niche overlap measure. These techniques have been
discussed in some detail by May (1975), Hanski (1978) and Slobodchikoff and Schulz (1980). May (1975) shows that the estimates of multidimensional niche overlap by products or sums of the one-dimensional values are special cases, the product estimate being valid when all species' resource dimensions are orthogonal, and the summation estimate when the resource dimensions are all linearly dependent. According to the empirical results of Hanski (1978), the product overlap yields better estimates than the summation estimate. Slobodchikoff and Schulz (1980), using the niche overlap measure designed by Pianka (1973), provide a test of dependence-independence of resource dimensions that may help one to choose between either the products or sums techniques for assessing the combined effects of the one-dimension component niche overlaps. The measures developed by Harner and Whitmore (1977) are multivariate methods, that assume a certain degree of multivariate normality and variance homogeneity and may be affected by the number of variables chosen. The measure that these authors prefer is their $\beta_1$, a multivariate measure of distributional overlap which is equivalent to the estimated rate of misclassification in multivariate discriminant analysis. The other measure is their multivariate version of the MacArthur-Levins measure.

Hypothesis testing

Although one can imagine many hypotheses to test in the context of niche overlap measurement, I will recognize two that I consider to be of most general interest. The first hypothesis asks whether the niches of two given species are identical or different, and the second asks whether an observed niche overlap is minimized by interspecific interactions in a nonrandom way.

First hypothesis: are the niches of two species identical or different?

As shown above, none of the available measures may help to test this hypothesis directly since none of them measures full niche overlap. The task of measuring multidimensional overlap would, at least, require the development of multivariate techniques permitting the estimation of density overlap in resource use with variables of very different nature.

Until such techniques are available to ecologists, the hypothesis must be limited to multivariate-normal and homogeneous variables (Harner and Whitmore 1977), or fragmented into hypotheses concerning the component niche overlaps, sensu Cody (1974) and May (1975).

Ludwig and Reynolds (1988) point out that the usefulness of niche overlap measures is limited, because they do not take resource availability into account and because no statistical tests are available. Several authors have also argued that if resource availability is not accounted for niche overlap measures are of limited value (e.g. Hurlbert 1978, Petraitis 1979). If niche overlap measures are taken as measures of competition intensity or estimates of competition coefficients, this may be true, since not only resource abundance levels, but also differences in resource productivity may be of importance (Lawlor 1980a). Moreover, when comparisons are made between communities, care should be taken that the comparisons are reasonable, since when the niche overlap measure is made on a resource state that reflects consumer-environment interactions such as prey category consumption, differences in relative availability of the resource categories between the communities will produce misleading overlap values (Lawlor 1980a). In such comparisons between communities, it is usually better to correct resource use values by the relative resource availabilities, as suggested by Zaret and Smith (1984): if $R_k$ denotes the availability of resource $k$ and $p_{ik}$ the relative use of resource $k$ by species $i$, then the overlap measure can be computed based on the proportion $p_{ij}/R_k$.

However, in both intra- and intercommunity comparisons, other sources of variability are to be considered, e.g. variability in resource use due to spatial heterogeneity in resource availability, heterogeneity in resource renewal rates and individual differences in resource selectivity. This means that replicated random sampling of resource use should be applied in order to sample the variability of niche overlap between both species under test. With replicated data, standard statistical techniques can be applied to overlap measures (Horn 1966). To test the hypothesis of niche identity between two species in a standard way, one can, for example, apply the following method. For a replicated series of $n$ niche overlap measures between the two species, there are $n$ resource use data sets for each species. This means that one can draw a defined number of possible intraspecific pairs within each species and calculate the same number of interspecific niche overlaps. This number of independent intraspecific niche overlap measures one can obtain is, for $k=2$:

$$\binom{n}{k} = \frac{n!}{k! (n-k)!}$$

From the complete set of intraspecific niche overlap values which one can obtain for the two species, one can then draw a random set of $n$ pairs. Thus, this simple operation yields a replicated sample of intraspecific niche overlap besides the replicated sample of interspecific niche overlap. To test the hypothesis of no difference between the niches of both species, it suffices to test for differences between the two sets of intra- and interspecific niche overlap values. Absence of a significant difference will indicate niche identity, and a one-tailed difference where the interspecific values are
significantly lower than the intraspecific values will indicate significant niche difference between the two species. The case where interspecific values are higher than intraspecific may be neglected. Such a difference will arise only when intraspecific replicates are taken from extremely different habitats and, thus, may be considered a consequence of sampling bias. One may also perform separate tests between the set of interspecific overlaps and the two sets of intraspecific overlaps, i.e. consider the intraspecific overlaps of the two species separately when comparing with interspecific overlaps.

Other statistical procedures have been proposed to test niche overlap measures, e.g. by Garratt and Steinhorst (1976), Petraitis (1979), Maurer (1982), Zaret and Smith (1984) and Mueller and Altenberg (1985).

The technique suggested by Garratt and Steinhorst (1976) defines confidence intervals and tests of significance for the $C_\lambda$ and the $R_0$ measures of Horn (1966) on the basis of a non-parametric permutation test. However, their technique leads to very broad confidence intervals and it is of limited value for empirical applications (Garratt and Steinhorst 1976).

The niche overlap measure $\Phi_{ij}$ developed by Petraitis (1979) is designed to test whether the utilization curve of a given species may have been randomly drawn from the utilization curve of another species. This method also permits one to test whether the overlap between species $i$ and $k$ is greater than the overlap between $i$ and a second species $m$. A general niche overlap measure ($G$) has been developed on the same basis (Petraitis 1979) and tests whether there is complete overlap between several different species. As pointed out by Petraitis, the limitation of his methods is that the relative strength of overlap is only tested against complete overlap (Petraitis 1979).

The contributions of Maurer (1982), Zaret and Smith (1984) and Mueller and Altenberg (1985) all furnish confidence interval estimators for several measures (Table 1). As Zaret and Smith (1984) underline, variance estimates and confidence intervals permit comparison of the observed overlap with any specified value, or comparison of different overlap values of the same species at several times or areas. While useful for any case where such tests are required, they are of limited use to test the hypothesis of niche identity, since a predetermined value such as complete overlap must be hypothesized.

**Second hypothesis: are observed niche overlap values minimized in a nonrandom way?**

This type of hypothesis addresses the problem of community organization. Can one test the possibility that interspecific interactions have minimized overlap? To test such an hypothesis one needs a neutral model that can predict the degree of overlap expected under random conditions (Pianka 1976). Several models of this type have been proposed (e.g. Sale 1974, Lawlor 1980b, Cole 1981, and Sugihara 1986). The applicability and validity of the conclusions one can draw from these models are sometimes limited by the assumptions of the model (Sale 1974). However, the recent shuffled-sticks model (Sugihara 1986), which is an analytical method to detect nonrandom niche displacement patterns that might result from competitive interactions, shows that there are promising possibilities to test the hypothesis of nonrandom niche overlap minimization in community organization. The shuffled-sticks model applies to a single niche dimension and calculates the probability that an observed overlap value is larger or equal to the random value expected with the observed parameters of resource availability and resource use. The model is supported by some empirical results (Cole 1981, Sugihara 1986, Neet pers preparation, S. Pimm pers comm.).

**Recommendations for empirical applications**

I considered several criteria for retaining the most useful niche overlap measures among those available, i.e. simplicity of computation, symmetry, low bias, availability of confidence interval estimators, and the possibility of performing comparisons between communities. Only a few measures combine all these qualities for single resource measures: the Freeman-Tukey statistic (Matusita 1955), the original measure of Morisita (1959), and the Euclidean distance measure proposed by Levins (1968). Among these, the Freeman-Tukey statistic is the simplest to compute. It has also been suggested as a niche breadth measure (Smith 1982), and its properties have been studied in detail (Smith 1982; Smith and Zaret 1982; Zaret and Smith 1984):

\[ FT_{ij} = \sum_{r=1}^{k} (p_{ir} - p_{jr})^{1/2} \]

where $k$ is the number of resource categories $r$, and $p_{ir}$ is the frequency of use of resource $r$ relative to the total amount of resource use. For reasonable sample sizes, and overlap values not too close to 1, this measure may be treated approximately as normal, and an approximation of 95% confidence interval (Smith 1982) is:

\[ \sin (\arcsin (FT_{ij}) + \frac{1.96}{2 \sqrt{N}}) \]

For the measurement of niche overlap in several resource states, the products and summation techniques of component overlaps may be used in empirical applications after the dependence-independence of resource states has been tested (Slobodchikoff and Schulz 1980). Alternatively, the multivariate method of Harner and Whitmore (1977) may be adopted under the li-
miting assumptions of multivariate normality and variance homogeneity, which may be tested with appropriate statistical techniques, e.g. Box’s M test.

When a more detailed approach for intracommunity studies is needed, the measures of Petravitis (1979, 1985) are useful for direct tests of complete niche identity of two or more species. Hurlbert’s measure (Hurlbert 1978) is recommended when, in addition to niche overlap, one is interested in similarity in intracommunity selectivity patterns. In such cases, one should also consider the possibility of computing selectivity overlap by applying a niche overlap measure to the selectivity values of Chesson (1978).

Conclusions

Available measures for single resource dimension are, at least, sufficient for empirical applications. They are numerous, and several statistical analyses are available. However, the data used with these measures should be carefully chosen and their relevance seriously considered since, in the case of trophic niche comparisons where prey categories are used, low ordinal levels of prey identification may be a source of considerable bias (Greene and Jaksic 1983) and, in the case of continuous variables, scale choice may also influence the outcome of the measures (Case 1984).

For multiple resource overlap measurement, technical foundations do exist, and, when distributional properties and dependence between variables are known, valuable results may be obtained. Nevertheless, the available methods are limited and progress is expected to occur in this domain, especially from new approaches to the measurement of overlap in multidimensional space (E. Feoli and P. Ganis, pers. comm.).

Hypothesis testing

As shown above, for several niche overlap measures, confidence interval estimators have been developed which permit tests of hypotheses such as equality between two overlap measures, and equality between an overlap measure and a fixed value. As pointed out by Horn (1966), in practical circumstances, standard statistical techniques applied to replicated simple measures may be of more value than exact tests based on the theoretical variation of more complex indexes. In the case of the hypothesis of niche identity between two species, within a community or between different communities, it seems that the procedure suggested by Horn (1966) and developed in this paper is, at present, the best alternative.

The present status of the niche overlap concept

This review has deliberately taken niche overlap as an ecological descriptor of similarity in resource use, in the sense of Abrams (1980). Much of the work that has been done on the organization of ecological communities makes explicit reference to competition and niche overlap (e.g. May 1973, Cohen 1978, Lawlor 1980b, Sugihara 1986). With respect to community organization, it makes sense to say that observed overlaps may have been influenced by interactions between species, in the present or in the past, i.e. without having to make any assumptions on the actual level of intensity of interspecific competition. However, when one says that competitors can tolerate a certain degree of niche overlap lower than a limiting similarity (Mac Arthur and Levins 1967, May 1973), one admits that high niche overlap will lead to a high degree of competition. On the other hand, the actual consequence of a high degree of competition may be a low degree of effective overlap, as predicted in the frame of habitat selection theory (e.g. Rosenzweig 1981, Pimm and Rosenzweig 1981).

It appears that the contradiction one may find between such statements is due to the fact that the niche overlap concept has itself more specific underlying concepts, a problem first approached by Lawlor (1980a) and clearly recognized by Schoener (1986). At present at least three different concepts may be recognized: the realized niche overlap, the similarity in use of given resources, and the value of interspecific competition coefficients. Our understanding of the consequences of the ecological similarity of species for community organization will certainly depend on a better separation of these concepts in the future.

Acknowledgements. I would like to thank Prof. J. Hauser and Prof. B.O. Bengtsson for stimulating discussions while I started working with niches, Dr. R. Arditii for advice on some mathematical problems, Dr. M.A. Burgman for comments on the manuscript and Prof. P. Vogel and the University of Lausanne for support.

REFERENCES


PETRAITIS, P.S. 1983. Presentation of niche measure relationships when more than three resource classes are involved. Ecology 64: 1318-1320.


VAN BELLE, G. and I. AHMAD. 1974. Measuring affinity of


Manuscript received: January 1989